

LUD 5538.1 CIP - JEL/NDH (09807399)

Claim 139: The isolated nucleic acid molecule of claim 110, comprising SEQ ID NO: 8.

Claim 140: The isolated nucleic acid molecule of claim 110, consisting of SEQ ID NO: 7.

Claim 141: The isolated nucleic acid molecule of claim 110, consisting of SEQ ID NO: 6.

Claim 142: The isolated nucleic acid molecule of claim 110, consisting of SEQ ID NO: 7.

Claim 143: The isolated nucleic acid molecule of claim 110, consisting of SEQ ID NO: 8.

**REMARKS**

The preceding amendment is presented in accordance with 37 CFR § 1.121(h). Claims 110-143 will be pending.

The amendment address points 3 and 10 of the office action. The objections to claims 108 and 109 are obviated via the sequence listing. Applicants note that the sequence listing objection was not sent with the June 5 office action, but was telefaxed on June 12, 2001.

Prior to addressing the bases for rejecting the claims, it is worthwhile to discuss what the empirical evidence provided in the application shows.

In example 1, a melanoma cell line is discussed. The prior art describes how various members of the CT family of antigens are expressed by this cell. The cell was used to determine expression of other members of the family. It should be borne in mind that the term "CT antigen" is discussed in the specification.

The cells were studied, using the "SEREX" methodology. This method is known in the art. To perform SEREX, total RNA is removed from cells, converted to cDNA and, a cDNA library is constructed in a vector. In this case, " $\lambda$ ZAP" was used. The  $\lambda$ ZAP cDNA library was then used to transform E. coli.

After E. coli are transformed, the transformed cells are used to screen samples for antibodies. Allogeneic samples were used in example 1. Essentially, one is determining if the

LUD 5538.1 CIP - JEL/NDH (09807399)

protein encoded by the cDNA used to transform the E. coli reacts with antibodies in the sample. One knows that the reaction is not with a "native" E. coli protein, because in the SEREX method, such antibodies are "stripped" prior to reaction with the transformants. Applicants refer the examiner to the references cited at page 5, i.e., Sahin, et al, Proc. Natl. Acad Sci USA 92:11810-11813, and U.S. Patent No. 5,698,396, both of which are incorporated by reference. Full details of the SEREX methodology are described in these references. Hence, one knows that, if there are any positives, the reactive protein was from the cDNA library.

The yield in this experiment was high, i.e., there were 65 positives.

Example 2 describes, simply, attempts to identify the positive materials. Standard methods were used.

Table 8, at page 8, shows that of the 65 positives 18 were known CT antigens. Hence, one knows from this that one can identify known members of the CT family. Ten sequences were not identifiable in public sequence libraries. There were also, 33 clones which belonged to the "KOC" family.

Examples 3 and 5 deal with CT antigen, i.e., "CT7." Please see US Patent No. 6,297,364.

Example 6 describes sequence analysis of the 33 members of the family found in the experiments of example 2. As this example makes manifestly clear, two different genes, referred to as KOC-2 and KOC-3 were identified. They were clearly not artifacts, because when a testicular cDNA library was analyzed, again using allogeneic samples, the molecules were found.

In example 7, the DNA expression pattern of the genes was studies. KOC-2 was found in normal tissue, but only testis and, with modification of the assay, in normal kidney cells. This is a "classic" CT expression pattern. As noted in the application, "CT" means "cancer/testis."

The KOC-3 gene was found to be expressed in both cancer and normal cells.

Now, the examiner has rejected all of claims 53-73, 80-84, 108 and 109, arguing that the claims are not supported by either "a specific substantial utility or a well-established utility."

The examiner states that "the specification fails to demonstrate a utility for the nucleic acids of SEQ ID NOS: 5-8, as the specification fails to correlated the presence of the protein encoded by SEQ ID NOS: 5-8 in clinical samples." The examiner then goes on to hypothesize that (i) antibodies can cross react with different proteins, and (ii) that the polynucleotides may

LUD 5538.1 CIP - JEL/NDH (09807399)

not be expressed due to any of a number of hypotheses. The examiner also states that "under conditions of recombinant expression, epitopes of KOC-2 and 3 which are not normally expressed in a patient within the context of MHC could be expressed by E. coli."

All of this is unproven speculation. The facts are the following. Something in the serum of patient NW38 reacted with a protein expressed by cDNA from an allogeneic source. Antibodies were present in the cancer patient's serum. They did not react with an E. coli protein because, as has been pointed out, supra, when SEREX is carried out, any reactivity with E. coli proteins is eliminated by use of controls. Again, the examiner is directed to the references which describe the SEREX methodology.

With respect to the elaborate argument regarding factors affecting translation of proteins, this is all well and good, but it is irrelevant. It is irrelevant because there were antibodies in the sample analyzed. Antibodies are produced in response to a protein. Hence, at some point, NW38 (the source of the sample), did have the protein in his or her blood. It got there by being expressed. The issue for CT antigens, as was set forth very clearly in the specification, is: are they expressed in individuals with cancer? The evidence indicates that KOC-2 and KOC-3 are so expressed, regardless of what happened to the irrelevant ferritin polypeptide, ornithine decarboxylase, and p glycoprotein discussed by the examiner, and not shown to be cancer antigens. With respect to the failure of protein and mRNA levels to correlate for p53, again, this is irrelevant. Applicants have not claimed a quantitative assay. The correct question is: is the molecule expressed in cancer patients and usable for diagnosis? The correlation of p53 with cancer is too well known to comment on.

With respect to the examiner's hypothetical argument that "antibodies bind to epitopes in proteins and antibodies can cross react with different proteins containing the same or similar proteins," this again ignores the evidence. In the SEREX methodology, antibodies against a host cell, e.g., E. coli, are removed in a stripping step. All that is left are antibodies which bind to non-E. coli proteins. Applicants used a cDNA library. It is well known that, when using a cDNA library, a single cDNA molecule is used as the transfectant. Hence, if an antibody reacts with a transfected or transformed cell, it must be reacting with the foreign protein, since antibodies to native proteins have been eliminated. Hence, all of the examiner's points are based upon generalizations which do not stand up to analysis when the data are considered. With respect to the statement bridging pages 4-5:

LUD 5538.1 CIP - JEL/NDH (09807399)

"The data indicating that the KOC-2 polynucleotide was restricted to the normal testis and that the KOC-3 polynucleotide was universal in normal tissues."

This is absolutely an incorrect reading of the data. First of all, there were antibodies against KOC-2 and KOC-3 found in the patient (NW38) sample. Hence, there was no immunotolerance. Second, the data state that "the only positive normal tissue was testis." These do not state that this was the only positive tissue. There is a very big difference. Further, if the examiner is going to argue that levels of expression can vary, then she must concede that, notwithstanding expression of KOC-3 in normal tissue, cancer tissues express it at a level such that an immune response is possible - especially since applicants have shown this.

Turning next to the rejection under 35 USC § 112, the statement of this rejection is not consistent. All claims are rejected, even though the examiner states that the claims are enabling for "polynucleotides comprising SEQ ID NOS: 5-8..." If this is the case, then claims 54-70 and 108 and 109 should not be rejected, as claims 54-70 expressly recite what is said to be enabled, and claims 108 and 109 have nothing to do with the issues raised in the rejection. The rejection recites points "A" and "B," and it is noted that only claims 53, 71-73 and 80-84 are rejected. They are no reasons given for rejecting the remaining claims.

Applicants contacted the examiner regarding this issue on October 22. The undersigned was advised that, if the utility rejection were overcome, then the issue at point 7 would be moot. It is noted that claims 54-70 are not rejected in the discussion following point 7. Claims 53 & 54 were combined to form claim 110. As such, claim 110 and claims dependent thereon cannot be rejected under 35 USC §112.

In view of the foregoing, allowance of this application is believed proper and is urged.

LUD 5538.1 CIP - JEL/NDH (09807399)

In view of the lack of clarity in this rejection, applicants cannot address it. The examiner is invited to clarify the rejection so it can be addressed. At such time, applicants will respond fully.

Respectfully submitted,

FULBRIGHT & JAWORSKI, L.L.P.



Norman D. Hanson, Esq.  
Registration No. 30,946

666 Fifth Avenue  
New York, New York 10103-3198  
Telephone: 212-318-3168  
Telecopier: 212-318-3400

<110> Chen, Yao-Tseng  
 Gure, Ali  
 Tsang, Solam  
 Stockert, Elisabeth  
 Jager, Elke  
 Knuth, Alexander  
 Old, Lloyd J.

<120> Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen, The  
 Antigens Per Se, And Uses Thereof

<130> LUD 5538.1

<140> 09/270,437

<141> 1999 - 03 - 16

<150> 09/061,709

<151> 1998-04-17

<160> 23

C2  
<210> 1  
<211> 4265  
<212> DNA  
<213> Homo sapiens  
<220>  
<400> 1

GTCTGAAGGA CCTGAGGCAT	TTTGTGACGA GGATCGTCTC	AGGTCAGCGG AGGGAGGGAGA	60
CTTATAGACC TATCCAGTCT	TCAAGGTGCT CCAGAAAGCA	GGAGTTGAAG ACCTGGGTGT	120
GAGGGACACA TACATCCTAA	AAGCACACAA GCAGAGGAGG	CCCAGGCAGT GCCAGGGAGTC	180
AAGGTTCCA GAAGACAAAC	CCCCTAGGAA GACAGGCGAC	CTGTGAGGCC CTAGAGCACC	240
ACCTTAAGAG AAGAACAGCT	GTAAGCCGGC CTTTGTAGA	GGCATCATGG GGGACAAGGA	300
TATGCCTACT GCTGGATGC	CGAGTCTCT CCAGAGTTCC	TCTGAGAGTC CTCAGAGTTG	360
TCCTGAGGGG GAGGGACTCC	AGTCTCCTCT CCAGATTCCC	CAGAGTTCTC CTGAGAGCGA	420
CGACACCTTG TATCCTCTCC	AGAGTCCTCA GAGTCGTTCT	GAGGGGGAGG ACTCCTCGGA	480
TCCTCTCCAG AGACCTCCTG	AGGGGAAGGA CTCCCAGTCT	CCTCTCCAGA TTCCCCAGAG	540
TTCTCCTGAG GGGCACGACA	CCCAGTCTCC TCTCCAGAAT	TCTCAGAGTT CTCCAGAGGG	600
GAAGGACTCC CTGTCTCCTC	TAGAGATTTC TCAGAGCCCT	CCTGAGGGTG AGGATGTCCA	660
GTCTCCTCTG CAGAATCCTG	CGAGTCCCTT CTTCCCTCT	GCTTTATTGA GTATTTCCA	720
GAGTCCCCCT GAGAGTATT	AAAGTCCTT TGAGGGTTT	CCCCAGTCTG TTCTCCAGAT	780
TCTGTGAGC CCCGCTCCT	CCTCCACTTT AGTGAAGTATT	TTCCAGAGTT CCCCTGAGAG	840
TACTCAAAGT CCTTTGAGG	GTTCCTTCCA GTCTCCACTC	CAGATTCCCTG TGAGCCGCTC	900
CTTCTCCTCC ACTTTATTGA	GTATTTTCCA GAGTCCCCCT	GAGAGAACCTC AGAGAACCTC	960
TGAGGGTTTT GCACAGTCTC	CTCTCCAGAT TCCTGTGAGC	TCCTCCTCGT CCTCCACTTT	1020
ACTGAGTCTT TTCAGAGTT	CCCCCTGAGAG AACTCAGAGT	ACTTTTGAGG GTTTTCCCCA	1080
GTCTCCACTC CAGATTCTG	TGAGCCGCTC CTTCTCCTCC	ACTTTATTGA GTATTTCCA	1140
GAGTTCCCCCT GAGAGAACCTC	AGAGTACTTT TGAGGGTTTT	CCCCAGTCTC CTCTCCAGAT	1200
TCCTGTGAGC CCCCTCTCT	CCTCCACTTT AGTGAAGTATT	TTCCAGAGTT CCCCTGAGAG	1260
AACTCAGAGT ACTTTGAGG	GTTCCTTCCA GTCTCCCTC	CAGATTCCCTG TGAGCTCCTC	1320
CTTCTCCTCC ACTTTATTGA	GTCTTTTCCA GAGTTCCCCCT	GAGAGAACCTC AGAGTACTTT	1380
TGAGGGTTTT CCCCAGTCTC	CTCTCCAGAT TCCTGGAAGC	CCCTCCTCT CCTCCACTTT	1440
ACTGAGTCTT TTCAGAGTT	CCCCCTGAGAG AACTCACAGT	ACTTTTGAGG GTTTTCCCCA	1500
GTCTCCTCTC CAGATTCTA	TGACCTCCTC CTTCTCCTCT	ACTTTATTGA GTATTTACA	1560
GAGITCTCCT GAGAGTGCTC	AAAGTGCTT TGAGGGTTTT	CCCCAGTCTC CTCTCCAGAT	1620
TCCTGTGAGC TCCTCTTTCT	CCTACACTT ATTGAGTCTT	TTCCAGAGTT CCCCTGAGAG	1680
AACTCACAGT ACTTTGAGG	GTTCCTTCCA GTCTCCCTC	CAGATTCCCTG TGAGCTCCTC	1740

C2  
 w  
 CTCCTCCTCC TCCACTTTAT TGAGTCTTT CCAGAGTTCC CCTGAGTGTA CTCAAAGTAC 1800  
 TTTGAGGGT TTTCCCAGT CCTCTCTCCA GATTCTCTAG AGTCCTCTG AAGGGGAGAA 1860  
 TACCCATTCT CCTCTCCAGA TTGTTCCAAG TCTTCCTGAG TGGGAGGACT CCTGTCTCC 1920  
 TCACTACTTT CCTCAGAGCC CCTCTCAGGG GGAGGACTCC CTATCTCCTC ACTACTTCC 1980  
 TCAGAGCCT CCTCAGGGGG AGGACTCCCT GTCTCCTCAC TACTTCTC AGAGCCTCA 2040  
 GGGGGAGGAC TCCCTGTCTC CTCACTACTT CCTCTCAGAGC CCTCCTCAGG GGGAGGACTC 2100  
 CATGTCTCCT CTCTACTTTCTC CCTAGAGTCC TCTTCAGGGG GAGGAATTCC AGTCTTCTCT 2160  
 CCAGAGCCT GTGAGCATCT GTCCTCCTC CACTCCATCC AGTCTTCCCC AGAGTTCCC 2220  
 TGAGAGTCT CAGAGTCCTC TGAGGGGGC TGTCAGTCT CCTCTCCATA GTCCTCAGAG 2280  
 CCTCTCTGAG GGGATGCACT CCCAATCTCC TCTCCAGAGT CCTGAGAGTG CCTCTGAGGG 2340  
 GGAGGATTCC CGTCTCCTC TCCAAATTCC TCAGAGTCTT CTGAGGGAG AGGACTCCCT 2400  
 GTCTCTCTC CATTTCCTC AGAGTCTCC TGAGTGGGAG GACTCCCTCT CCTCTCTCCA 2460  
 CTTTCCTCAG TTTCCTCTC AGGGGGAGGA CTTCCAGTCT TCTCTCCAGA GTCCTGTGAG 2520  
 TATCTGCTCC TCCTCCACTT CTTGAGTCT TCCCCAGAGT TTCCCTGAGA GTCCTCAGAG 2580  
 TCCTCTGAG GGGCTGCTC AGTCTCTCTC CCAGAGACCT GTCAGCTCTC TCTCTCCTA 2640  
 CACTTAGCG AGTCTTCTCC AAAGTTCCA TGAGAGTCT CAGAGTCCT CTGAGGGGCC 2700  
 TGCCCAGTCT CCTCTCCAGA GTCCTGTGAG CTCCCTCCCC TCCTCCACTT CATCGAGTCT 2760  
 TCCCCAGAGT TCTCTGTGA GTCCTCTCC CTCCCTCCACT TCATCGAGTC TITCCAAGAG 2820  
 TCCCCCTGAG AGTCTCTCC AGAGTCTGT GATCTCTTC TCCTCCTCCA TTCAITGAG 2880  
 CCCATTCACT GAAGAGTCCA GCAGCCCCAGT AGATGAATAT ACAAGTTCTC CAGACACCTT 2940  
 GCTAGAGAGT GATTCTTGA CAGACAGCGA GTCCTGTATA GAGAGCGAGC CCTTGTTCAC 3000  
 TTATACACTG GATGAAAAGG TGGACGAGTT GGCAGGGTTT CTTCTCCTCA AATATCAAGT 3060  
 GAAGCAGCCT ATCACAAAGG CAGAGATGTC GACGAATGTC ATCAGCAGGT ACACGGCTA 3120  
 CTTTCTGTG ATCTCAGGA AAGCCCTGTA GTTCATAGAG ATACTTTTG GCATTTCCT 3180  
 GAGAGAAGTG GACCTGTATG ACTCCTATGT CTTGTAAAC ACATTAGACC TCACCTCTGA 3240  
 GGGGTGCTG AGTGTATGAGC AGGGCATGTC CCAGAACCGC CTCCCTGATT TCATTCTGAG 3300  
 TATCATCTC ATAAAGGGCA CCTATGCCCTC TGAGGAGGTC ATCTGGGATG TGCTGAGTGG 3360  
 AATAGGGGTG CGTGTGGGA GGGAGCAGTT TGCCCTTGGG GAGCCCAGGG AGCTCCTCAC 3420  
 TAAAGTTGG GTGCAAGAAC ATTACCTAGA GTACCGGGAG GTGCCCAACT CTTCTCCTCC 3480  
 TCGTTACGAA TTCTGTGGG GTCCAAGAGC TCATTCAAGAA GTCATTAAGA GGAAAGTAGT 3540  
 AGAGTTTTG GCCATGCTAA AGAATACCGT CCCTATTAC TTTCCATCCT CTTACAAGGA 3600  
 TGCTTTGAAA GATGTGGAAG AGAGAGCCCA GGCCATAATT GACACCACAG ATGATTCGAC 3660  
 TGCCACAGAA AGTCAAGCT CCAGTGTAT GTCCCCCAGC TTCTCTCTG AGTGAAGTCT 3720  
 AGGGCAGATT CTTCCCTCTG AGTTGAAGG GGGCAGTCGA GTTCTACGT GGTGGAGGGC 3780  
 CTGGTGTGAGG CTGGAGAGAA CACAGTGTCA TTTGCATTTC GTTCCATAT GGGTAGTTAT 3840  
 GGGGTTTACCT GTGTTTACTT TTGGGTATTT TTCAAATGCT TTCTCTATTA ATAACAGGTT 3900  
 TAAATAGCTT CAGATCCTA GTTATGCAATGCACT ATGAGTCGCA CATGTATTGC TGTTTTCTG 3960  
 GTTTAAGAGT AACAGTTGA TATTTGTAA AAACAAAAAC ACACCCAAAC ACACCACATT 4020  
 GGGAAAACCT TCTGCCTCAT TTTGTGATGT GTCACAGGTT AATGTGGTGT TACTGTAGGA 4080  
 ATTTTCTGAA AACTGTGAAG GAACTCTGCA GTTAAATAGT GGAATAAAAGT AAAGGATTGT 4140  
 TAATGTTGCA ATTCCTCAG GTCTTTAGT CTGTTGTTCT TGAAAACCTAA AGATACATAC 4200  
 CTGGTTTGCT TGGCTTACGT AAGAAAGTCG AAGAAAGTAA ACTGTAATAA ATAAAAGTGT 4260  
 CAGTG 4265

<210> 2  
 <211> 1142  
 <212> PRT  
 <213> Homo sapiens  
 <220>  
 <400> 2

Met Gly Asp Lys Asp Met Pro Thr Ala Gly Met Pro Ser Leu Leu Gln  
                       5                    10                    15  
 Ser Ser Ser Glu Ser Pro Gln Ser Cys Pro Glu Gly Glu Asp Ser Gln  
                       20                    25                    30  
 Ser Pro Leu Gln Ile Pro Gln Ser Ser Pro Glu Ser Asp Asp Thr Leu  
                       35                    40                    45  
 Tyr Pro Leu Gln Ser Pro Gln Ser Arg Ser Glu Gly Glu Asp Ser Ser  
                       50                    55                    60  
 Asp Pro Leu Gln Arg Pro Pro Glu Gly Lys Asp Ser Gln Ser Pro Leu

	65	70	75	80
Gln Ile Pro Gln Ser Ser Pro Glu Gly Asp Asp Thr Gln Ser Pro Leu				
85		90		95
Gln Asn Ser Gln Ser Ser Pro Glu Gly Lys Asp Ser Leu Ser Pro Leu				
100	105		110	
Glu Ile Ser Gln Ser Pro Pro Glu Gly Glu Asp Val Gln Ser Pro Leu				
115	120		125	
Gln Asn Pro Ala Ser Ser Phe Phe Ser Ser Ala Leu Leu Ser Ile Phe				
130	135		140	
Gln Ser Ser Pro Glu Ser Ile Gln Ser Pro Phe Glu Gly Phe Pro Gln				
145	150		155	160
Ser Val Leu Gln Ile Pro Val Ser Ala Ala Ser Ser Ser Thr Leu Val				
165		170		175
Ser Ile Phe Gln Ser Ser Pro Glu Ser Thr Gln Ser Pro Phe Glu Gly				
180		185		190
Phe Pro Gln Ser Pro Leu Gln Ile Pro Val Ser Arg Ser Phe Ser Ser				
195		200		205
Thr Leu Leu Ser Ile Phe Gln Ser Ser Pro Glu Arg Ser Gln Arg Thr				
210		215		220
Ser Glu Gly Phe Ala Gln Ser Pro Leu Gln Ile Pro Val Ser Ser Ser				
225		230		235
Ser Ser Ser Thr Leu Leu Ser Leu Phe Gln Ser Ser Pro Glu Arg Thr				
245		250		255
Gln Ser Thr Phe Glu Gly Phe Pro Gln Ser Pro Leu Gln Ile Pro Val				
260		265		270
Ser Arg Ser Phe Ser Ser Thr Leu Leu Ser Ile Phe Gln Ser Ser Pro				
275		280		285
Glu Arg Thr Gln Ser Thr Phe Glu Gly Phe Ala Gln Ser Pro Leu Gln				
290		295		300
Ile Pro Val Ser Pro Ser Phe Ser Ser Thr Leu Val Ser Ile Phe Gln				
305		310		315
Ser Ser Pro Glu Arg Thr Gln Ser Thr Phe Glu Gly Phe Pro Gln Ser				
325		330		335
Pro Leu Gln Ile Pro Val Ser Ser Phe Ser Ser Thr Leu Leu Ser				
340		345		350
Leu Phe Gln Ser Ser Pro Glu Arg Thr Gln Ser Thr Phe Glu Gly Phe				
355		360		365
Pro Gln Ser Pro Leu Gln Ile Pro Gly Ser Pro Ser Phe Ser Ser Thr				
370		375		380
Leu Leu Ser Leu Phe Gln Ser Ser Pro Glu Arg Thr His Ser Thr Phe				
385		390		395
Glu Gly Phe Pro Gln Ser Pro Leu Gln Ile Pro Met Thr Ser Ser Phe				
405		410		415
Ser Ser Thr Leu Leu Ser Ile Leu Gln Ser Ser Pro Glu Ser Ala Gln				
420		425		430
Ser Ala Phe Glu Gly Phe Pro Gln Ser Pro Leu Gln Ile Pro Val Ser				
435		440		445
Ser Ser Phe Ser Tyr Thr Leu Leu Ser Leu Phe Gln Ser Ser Pro Glu				
450		455		460
Arg Thr His Ser Thr Phe Glu Gly Phe Pro Gln Ser Pro Leu Gln Ile				
465		470		475
Pro Val Ser Ser Ser Ser Ser Ser Thr Leu Leu Ser Leu Phe Gln				
485		490		495
Ser Ser Pro Glu Cys Thr Gln Ser Thr Phe Glu Gly Phe Pro Gln Ser				
500		505		510
Pro Leu Gln Ile Pro Gln Ser Pro Pro Glu Gly Glu Asn Thr His Ser				
515		520		525
Pro Leu Gln Ile Val Pro Ser Leu Pro Glu Trp Glu Asp Ser Leu Ser				
530		535		540

Pro His Tyr Phe Pro Gln Ser Pro Pro Gln Gly Glu Asp Ser Leu Ser  
 545 550 555 560  
 Pro His Tyr Phe Pro Gln Ser Pro Pro Gln Gly Glu Asp Ser Leu Ser  
 565 570 575  
 Pro His Tyr Phe Pro Gln Ser Pro Gln Gly Glu Asp Ser Leu Ser Pro  
 580 585 590  
 His Tyr Phe Pro Gln Ser Pro Pro Gln Gly Glu Asp Ser Met Ser Pro  
 595 600 605  
 Leu Tyr Phe Pro Gln Ser Pro Leu Gln Gly Glu Phe Gln Ser Ser  
 610 615 620  
 Leu Gln Ser Pro Val Ser Ile Cys Ser Ser Ser Thr Pro Ser Ser Leu  
 625 630 635 640  
 Pro Gln Ser Phe Pro Glu Ser Ser Gln Ser Pro Pro Glu Gly Pro Val  
 645 650 655  
 Gln Ser Pro Leu His Ser Pro Gln Ser Pro Pro Glu Gly Met His Ser  
 660 665 670  
 Gln Ser Pro Leu Gln Ser Pro Glu Ser Ala Pro Glu Gly Glu Asp Ser  
 675 680 685  
 Leu Ser Pro Leu Gln Ile Pro Gln Ser Pro Leu Glu Gly Glu Asp Ser  
 690 695 700  
 Leu Ser Ser Leu His Phe Pro Gln Ser Pro Pro Glu Trp Glu Asp Ser  
 705 710 715 720  
 Leu Ser Pro Leu His Phe Pro Gln Phe Pro Pro Gln Gly Glu Asp Phe  
 725 730 735  
 Gln Ser Ser Leu Gln Ser Pro Val Ser Ile Cys Ser Ser Ser Thr Ser  
 740 745 750  
 Leu Ser Leu Pro Gln Ser Phe Pro Glu Ser Pro Gln Ser Pro Pro Glu  
 755 760 765  
 Gly Pro Ala Gln Ser Pro Leu Gln Arg Pro Val Ser Ser Phe Phe Ser  
 770 775 780  
 Tyr Thr Leu Ala Ser Leu Leu Gln Ser Ser His Glu Ser Pro Gln Ser  
 785 790 795 800  
 Pro Pro Glu Gly Pro Ala Gln Ser Pro Leu Gln Ser Pro Val Ser Ser  
 805 810 815  
 Phe Pro Ser Ser Thr Ser Ser Leu Ser Gln Ser Ser Pro Val Ser  
 820 825 830  
 Ser Phe Pro Ser Ser Thr Ser Ser Leu Ser Lys Ser Ser Pro Glu  
 835 840 845  
 Ser Pro Leu Gln Ser Pro Val Ile Ser Phe Ser Ser Ser Thr Ser Leu  
 850 855 860  
 Ser Pro Phe Ser Glu Glu Ser Ser Ser Pro Val Asp Glu Tyr Thr Ser  
 865 870 875 880  
 Ser Ser Asp Thr Leu Leu Glu Ser Asp Ser Leu Thr Asp Ser Glu Ser  
 885 890 895  
 Leu Ile Glu Ser Glu Pro Leu Phe Thr Tyr Thr Leu Asp Glu Lys Val  
 900 905 910  
 Asp Glu Leu Ala Arg Phe Leu Leu Lys Tyr Gln Val Lys Gln Pro  
 915 920 925  
 Ile Thr Lys Ala Glu Met Leu Thr Asn Val Ile Ser Arg Tyr Thr Gly  
 930 935 940  
 Tyr Phe Pro Val Ile Phe Arg Lys Ala Arg Glu Phe Ile Glu Ile Leu  
 945 950 955 960  
 Phe Gly Ile Ser Leu Arg Glu Val Asp Pro Asp Asp Ser Tyr Val Phe  
 965 970 975  
 Val Asn Thr Leu Asp Leu Thr Ser Glu Gly Cys Leu Ser Asp Glu Gln  
 980 985 990  
 Gly Met Ser Gln Asn Arg Leu Leu Ile Leu Ser Ile Ile Phe  
 995 1000 1005  
 Ile Lys Gly Thr Tyr Ala Ser Glu Glu Val Ile Trp Asp Val Leu Ser  
 1010 1015 1020

Gly Ile Gly Val Arg Ala Gly Arg Glu His Phe Ala Phe Gly Glu Pro  
 1025 1030 1035 1040  
 Arg Glu Leu Leu Thr Lys Val Trp Val Gln Glu His Tyr Leu Glu Tyr  
 1045 1050 1055  
 Arg Glu Val Pro Asn Ser Ser Pro Pro Arg Tyr Glu Phe Leu Trp Gly  
 1060 1065 1070  
 Pro Arg Ala His Ser Glu Val Ile Lys Arg Lys Val Val Glu Phe Leu  
 1075 1080 1085  
 Ala Met Leu Lys Asn Thr Val Pro Ile Thr Phe Pro Ser Ser Tyr Lys  
 1090 1095 1100  
 Asp Ala Leu Lys Asp Val Glu Glu Arg Ala Gln Ala Ile Ile Asp Thr  
 1105 1110 1115 1120  
 Thr Asp Asp Ser Thr Ala Thr Glu Ser Ala Ser Ser Ser Val Met Ser  
 1125 1130 1135  
 Pro Ser Phe Ser Ser Glu  
 1140

<210> 3  
<211> 7  
<212> PRT  
<213> Homo sapiens  
<220>  
<400> 3

Pro Gln Ser Pro Leu Gln Ile  
 1 5

C2  
4  
<210> 4  
<211> 4159  
<212> DNA  
<213> Homo sapiens  
<220>  
<400> 4

GGTGGATGCG TTTGGGTGTT AGCTAGGCTT TTCTTTTCTT TTCTCTTTA AAACACATCT 60  
 AGACAAAGGAA AAAACAAGCC TCGGATCTGA TTTTCACTC CTCGTTCTG TGCTGGTTC 120  
 TTACTGTGTT TGTGTATTTT AAAGGCAGAGA AGACGAGGGG AACAAAACCA GCTGGATCCA 180  
 TCCATCACCG TGGGTGGTTT TAATTTTCG TTTTTCTCG TTATTTTTT TTAAACAACC 240  
 ACTCTTCACA ATGAACAAAC TGTATATCGG AAACCTCAGC GAGAACGCCG CCCCTCGGA 300  
 CCTAGAAAGT ATCTTCAGG ACGCCAAGAT CCCGGTGTG GGACCTTCC TGGTGAAGAC 360  
 TGGCTACGCG TTCGTGGACT GCCCGGACGA GAGCTGGGCC CTCAAGGCCA TCGAGGCGCT 420  
 TTCAGGTAAA ATAGAAACTGC ACGGGAAACC CATAGAAGTT GAGCACTCGG TCCCCAAAAG 480  
 GCAAAGGATT CGGAAACTTC AGATAAGAAA TATCCCGCCT CATTTCACAGT GGGAGGTGCT 540  
 GGATAGTTA CTAGTCCAGT ATGGAGTGGT GGAGAGCTGT GAGCAAGTGA ACAGTGACTC 600  
 GGAAACTGCA GTTGTAAATG TAACCTATTG CAGTAAGGAC CAAGCTAGAC AAGCACTAGA 660  
 CAAACTGAAT GGATTTCACT TAGAGAATT CACCTTGAAA GTAGCCTATA TCCCTGATGA 720  
 AATGGCCGCC CAGCAAAACC CCTTGAGCA GCCCGGAGGT CGCCGGGGC TTGGGCAGAG 780  
 GGGCTCCTCA AGGCAGGGGT CTCCAGGATC CGTATCCAAG CAGAAACCAT GTGATTTGCC 840  
 TCTGGCCCTG CTGGTTCCCA CCCAATTGTT TGGAGCCATC ATAGGAAAG AAGGTGCCAC 900  
 CATTGGAAC ATCACCAAAAC AGACCCAGTC TAAATCGAT GTCCACCGTA AAGAAAATGC 960  
 GGGGGCTGCT GAGAAGTCGA TTACTATCCT CTCTACTCCT GAAGGCACCT CTGCGGCTTG 1020  
 TAAGTCTATT CTGGAGGATTA TGCAATAAGGA AGCTCAAGAT ATAAAATCA CAGAAGAGAT 1080  
 CCCCTTGAG ATTCTAGCTC ATAATAACTT TGTGGACGT CTTATTGTT AAGAAGGAAG 1140  
 AAATCTTAAA AAAATTGAGC AAGACACAGA CACTAAAATC ACGATACTC CATTCAGGA 1200  
 ATTGACGCTG TATAATCCAG AACGCACTAT TACAGTTAAA GGCAATGTT AGACATGTGC 1260  
 CAAAGCTGAG GAGGAGATCA TGAAGAAAAT CAGGGAGTCT TATGAAAATG ATATTGCTTC 1320  
 TATGAATCTT CAAGCACATT TAATTCCCTGG ATTAATCTG AACGCCTTGG GTCTGTTCCC 1380

ACCCACTTCA GGGATGCCAC CTCCCACCTC AGGGCCCCCT TCAGCCATGA CTCCTCCCTA 1440  
 CCCGCAGTTT GAGCAATCAG AAACGGAGAC TGTTCATCAG TTATCCCCAG CTCTATCAGT 1500  
 CGGTGCCATC ATCGGCARGC AGGGCCAGCA CATCAAGCAG CTTCTCGCT TTGCTGGAGC 1560  
 TTCATTAAAG ATTGGTCCAG CGGAAGCACC AGATGCTAA GTGAGGATGG TGATTATCAC 1620  
 TGGACCACCA GAGGCTCAGT TCAAGGCTCA GGGAAAGAATT TATGGAAAAA TAAAGAAGA 1680  
 AACTTTGTT AGTCCTAAAG AAGAGGTGAA ACTTGAGCT CATATCAGAG TGCCATCCTT 1740  
 TGCTGCTGGC AGACTTATTG GAAAGAGGG CAAACGGTG AATGAACTTC AGAATTGTC 1800  
 AAGTGCAGAA GTTGTGTCC CTCGTGACCA GACACCTGAT GAGAATGACC AAGTGGTTGT 1860  
 CAAAATAACT GGTCACTTCT ATGCTTGCCTA GGTTGCCAG AGAAAAATTG AGGAAATTCT 1920  
 GACTCAGGTA AAGCAGCACC ACAACAGAA GGCTTGCAA AGTGGACCAC CTCAGTCAG 1980  
 ACAGGAAGTAA AGGCTCAGGA ACAGCCAC CACAGAGGA GATGCCAAC CAAAGACAGA 2040  
 TTGCTTAAAC AACAGATGGG CGCTGACCCC CTATCCAGAA TCACATGCAC AAGTTTTAC 2100  
 CTAGCCAGTT GTTCTGAGG ACCAGGCAAC TTTTGAACCTC CTGCTCTGT GAGAATGTAT 2160  
 ACTTTATGCT CTCTGAAATG TATGACACCC AGCTTTAAA CAAACAAACA ARCAAACAAA 2220  
 AAAAGGGTGG GGGAGGGAGG GAAAGAGAAG AGCTCTGCAC TTCCCTTTGT TGTAGTCTA 2280  
 CAGTATAACA GATATTCTAA TTCTTCTTAA TATTCCCCA TAATGCCAGA AATTGGCTTA 2340  
 ATGATGCTT CACTAAATTG ATCAAATAGA TTGCTCCTAA ATCCAATTGT TAAATTGGA 2400  
 TCAGAATAAT TATCACAGGA ACTTAAATGT TAAGCCATTA GCATAGAAAA ACTGTTCTA 2460  
 GTTTTATTT TACCTAACAC TAACATGAGT AACCTAAGGG AAGTGTGAA TGGTGTGGC 2520  
 AGGGGTATTA AACGTGCATT TTTACTCAAC TACCTCAGGT ATTCACTAAT ACAATGAAAA 2580  
 GCAAAATTGT TCCTTTTTT TGAAAATTAT ATATACTTTA TAATGATAGA AGTCCAACCG 2640  
 TTTTTAAAT AATAAATTAA AAATTAAACA GCAATCAGCT AACAGGCAA TTAAGATTCT 2700  
 TACTCTGGC TGGTGCAGT AAAGCTGGAA ATTAATTTC AGGGTTTTT GAGGCTTTG 2760  
 ACACAGTAT TAGTAAATC AAATGTTCAA AAATACGGAG CAGTGCCTAG TATCTGGAGA 2820  
 GCAGCACTAC CATTATTCTT TTCATTTATA GTTGGGAAAG TTTTGACGG TACTAACAAA 2880  
 GTGGTCGCAG GAGATTTGG AACGGCTGGT TAAATGGCT TCAGGAGACT TCAGTTTTT 2940  
 GTTTAGCTAC ATGATTGAAT GCATAATAAA TGCTTGTGC TTCTGACTAT CAATACCTAA 3000  
 AGAAAAGTGA TCAGTGAAGA GATGCAAGAC TTTCAACTGA CTGGCAAAA GCAAGCTTTA 3060  
 GCTTGTCTTA TAGGATGCTT AGTTGCCAC TACACTTCAG ACCAATGGGA CAGTCATAGA 3120  
 TGGTGTGACA GTGTTAAC GCAACAAAG GCTACATTTC CATGGGGCCA GCACTGTCT 3180  
 GAGCCTCACT AAGCTATTT GAAGATTTT AAGCACTGAT AAATTAAAAA AAAAAAAA 3240  
 AAATTAGACT CCACCTTAAG TAGTAAAGTA TAACAGGATT TCTGTATACT GTGCAATCAG 3300  
 TTCTTGAAA AAAAGTCAA AAGATAGAGA ATACAAGAAA AGTTTNGGG ATATAATTG 3360  
 AATGACTGT AAAACATATG ACCTTTGATA ACGAACTCAT TTGCTCACTC CTTGACAGCA 3420  
 AAGCCCAGTA CGTACAATTG TGTGGGGTGT GGGGGTCTC CAAGGCCACG CTGCTCTCTG 3480  
 AATTGATTT TTGAGTTTG GNTTGNAGA TGATCACAGN CATGTTACAC TGATCTTNA 3540  
 GGACATATNT TATAACCCCT TAAAAAAA ATCCOCTGCC TCATTCTTAT TTCGAGATGA 3600  
 ATTCGATAC AGACTAGATG TCTTCTGAA GATCAATTAG ACATTNTGAA AATGATTAA 3660  
 AGTGTTCCTTCTTAAATGTTCT CTGAAAACAA GTTCTTTTG TAGTTTAAC CAAAAAAAGTG 3720  
 CCTCTTTGTC CACTGGTTTC TCCTAGCATT CATGATTTT TTTTCACACA ATGAATTAAA 3780  
 ATTCGCTAAA TCATGGACTG GCTTCTGGT TGGATTCAG GTAAGATGTG TTTAAGGCCA 3840  
 GAGCTTTCT CAGTATTGAA TTTTTTCCC CAATATTGA TTTTTTAAA ATATAACACAT 3900  
 AGGAGCTGCA TTAAACACCT GCTGGTTAA ATTCTGTCTC ATTTCACTTC TAGCCTTTA 3960  
 GTATGCCNAAC TCAAAATTAA CTTTACTTA AGCATTTGTA ATTTGGAGTA TCTGGTACTA 4020  
 GCTAAGAAAT AATTNCATAA TTGAGTTTG TACTCNCCAA ANATGGGTCA TTCCTCATGN 4080  
 ATAATGTNCC CCCAATGCAG CTTCAATTTC CAGANACCTT GACCGCAGGAT AAATTTTTC 4140  
 ATCATTAGG TCCCCAAA 4159

<210> 5  
 <211> 1708  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <400> 5

AGGGACGGCTG	CCGCACCGCC	CCAGTTTAC	CCGGGGAGCC	ATCATGAAGC	TGAATGGCCA	60
CCAGTTGGAG	AACCATGCC	TGAAGGTCTC	CTACATCCCC	GATGAGCAGA	TAGCACAGGG	120
ACCTGAGAAT	GGGCGCCGAG	GGGGCTTTGG	CTCTCGGGGT	CAGCCCCGCC	AGGGCTCACC	180
TGTGGCAGCG	GGGGCCCCAG	CCAAGCAGCA	GCAAGTGGAC	ATCCCCCTTC	GGCTCCTGGT	240

GCCCACCCAG TATGTGGGTG CCATTATTGG CAAGGGAGGGG GCCACCATCC GCAACATCAC 300  
 AAAACAGACC CAGTCCAAGA TAGACGTGCA TAGGAAGGGG AACGCAGGTG CAGCTGAAAA 360  
 AGCCATCAGT GTGCACTCCA CCCCTGAGGG CTGCTCCTCC GCTGTAAAGA TGATCTTGG 420  
 GATTATGCAT AAAGAGGCTA AGGACACCAA AACGGCTGAC GAGGTTCCCC TGAAGATCCT 480  
 GCCCCATAAT AACTTTGTAG GGCGTCTCAT TGGCAAGGAA GGACGGAACC TGAAGAAGGT 540  
 AGAGCAAGAT ACCGAGACAA AAATACCATC CTCCTCGTT CAAGACCTTA CCCTTACAA 600  
 CCCTGAGAGG ACCATCACTG TGAAAGGGGG CATCGAGAAT TGTGCGAGGG CCGAGCAGGA 660  
 ATAATGAAG AAAGTTCGGG AGGCCTATGA GAATGATGTG GCTGCCATGA GCTCTCACCT 720  
 GATCCCTGGC CTGAACCTGG CTGCTGTAGG TCTTTTCCC GCTTCATCCA GCGCAGTCCC 780  
 GCCGCCTCCC AGCAGCGTTA CTGGGGCTGC TCCCTATAGC TCCTTTATGC AGGCTCCCGA 840  
 GCAGGAGATG GTGCGAGGTGT TTATCCCCGC CCAGGCAGTG GGCGCCATCA TCGGCAAGAA 900  
 GGGGCAGCAC ATCAAACAGC TCTCCCGTT TGCCAGCGCC TCCATCAAGA TTGCACCAACC 960  
 CGAACACCT GACTCCAAG TTCGTATGGT TATCATCACT GGACCGCCAG AGGCCCAATT 1020  
 CAAGGCTCAG GGAAGAATCT ATGGCAAAC CAAGGAGGAG AACTTCTTTG GTCCCAAGGA 1080  
 GGAAGTGAAG CTGGAGACCC ACATACGTGT GCCAGCATCA GCAGCTGGCC GGGTCATTGG 1140  
 CAAAGTGGG AAAACGGTGA ACGAGTTGCA GAATTTGACG GCAGCTGAGG TGGTAGTACC 1200  
 AAGAGACCAAG ACCCTGATG AGAACGACCA GGTCACTCGTGA AAAATCATCG GACATTCTA 1260  
 TGCCAGTCAG ATGGCTCAAC GGAAGATCCG AGACATCCTG GCCCAGGTTA AGCAGCAGCA 1320  
 TCAGAAGGGG CAGAGTAACC AGGCCCAGGC ACGGAGGAAG TGACCGAGCC CTCCCTGTCC 1380  
 CTTNGAGTCC AGGACAACAA CGGGCAGAAA TCGAGACTGT GCTCTCCCCG GCAGGCCTGA 1440  
 GAATGAGTGG GAATOCGGGA CACNTGGGC GGGCTGTAGA TCAAGTTTGC CCACTTGATT 1500  
 SAGAAAGATG TTCCAGTGG GAACCCCTGAT CTNTCAGCCC CAAACACCCCA CCCAATGGC 1560  
 CCAACACTGT NTGCCCTCG GGGTGTCAAGA AATTNTAGCG CAAGGCACTT TTAAACGTGG 1620  
 ATTGTTAAA GAAGCTCTCC AGGCCCCAACC AAGAGGGTGG ATCACACCTC AGTGGGAAGA 1680  
 AAAATAAAAT TTCCCTCAGG TTTTAAAAA 1708

<210> 6  
 <211> 3412  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <400> 6

C2  
 4  
 GGCAGCGGAG GAGGCGAGGA GCGCCGGGT A CCGGGCCGGG GGAGCCGGGG GCTCTCGGGG 60  
 AAGAGACGGA TGATGAACAA CCTTACATC GGGAACCTGA GCCCCCGCCGT CACCGCCGAC 120  
 GACCTCCGGC AGCTCTTGG GGACAGGAAG CTGCCCCCTGG CGGGACAGGT CCTGCTGAAG 180  
 TCCGGCTACG CCTCGTGG A CTACCCCGAC CAGAACTGGG CCATCCCGGC CATCGAGACC 240  
 CTCTCGGGTA AAGTGGAAAT GCATGGAAA ATCATGGAAG TTGATTACTC AGTCTCTAAA 300  
 AAGCTAAGGA GCAGGAAAT TCAGATTGCA AACATCCCTC CTCACCTGCA GTGGGAGGTG 360  
 TTGGATGGAC TTTGGCTCA ATATGGGACA GTGGAGAATG TGGAAACAAGT CAACACAGAC 420  
 ACAGAAACCG CGGTGTCAA CGTCACATAT GCAACAAGAG AAGAAGCAAA AATAGCCATG 480  
 GAGAAGCTAA GCGGGCATCA GTTGGAGAAC TACTCCTTCA AGATTTCTCA CATCCCGGAT 540  
 GAGAGGGTGA GCTCCCCCTC GCCCCCTCAG CGAGGCCAGC GTGGGGACCA CTCTTCCCGG 600  
 GAGCAAGGCC AGGCCCCCTGG GGGCACTTCT CAGGCCAGAC AGATTGATT CCCGCTGCCG 660  
 ATCCCTGGTCC CCACCCAGTT TGTGGTGCC ATCATCGGA AGGAGGGCTT GACCATAAAG 720  
 AACATCACTA AGCAGACCA GTCCCGGGTA GATATCCATA GAAAAGAGAA CTCTGGAGCT 780  
 GCAGAGAAGC CTGTCACCAT CCATGCCACC CGAGAGGGGA CTTCTGAAGC ATGCCGCATG 840  
 ATTCTTGGAAA TCATGCAAGA AGAGGCAGAT GAGACCAAAAC TAGCCGAAGA GATTCCCTTG 900  
 AAAATCTGG CACACAATGG CTTGGTTGG AAGACTGATTG GAAAAGAAGG CAGAAATTG 960  
 AAGAAAATTG AACATGAAAC AGGGACCAAG ATAACAATCT CATCTTGC A GGATTGAGC 1020  
 ATATACAACC CGGAAAGAAC CATCACTGTG AAGGGCACAG TTGAGGCCCTG TGCCAGTGCT 1080  
 GAGATAGAGA TTATGAAGAA GCTGCGTGAG GCCTTTGAAA ATGATATGCT GGCTGTTAAC 1140  
 CAACAAGCCA ATCTGATCCC AGGGTTGAAC CTCAGCGCAC TTGGCATCTT TTCAACAGGA 1200  
 CTGCTCCGTG TATCTCCACC AGCAGGGCCC CGGGGAGCTC CCCCCGCTGC CCCCTACAC 1260  
 CCCTTCAGTA CCCCACCTCGG ATACTCTCC AGCCTGTACC CCCATCACCA GTTGGCCCG 1320  
 TTCCCGCATC ATCACTCTTA TCCAGAGCAG GAGATTGTGA ATCTCTTCA CCAACCCAG 1380  
 GCTGTGGCG CCATCATCGG GAAGAAGGGG GCACACATCA AACAGCTGGC GAGATTGAGC 1440  
 GGAGCCCTCTA TCAAGATTGC CCCTGCGGAA GGCCAGACG TCAGCGAAAG GATGGTCATC 1500  
 ATCACCGGGC CACCGGAAGC CCAGTTCAAG CCCAGGGAC GGATCTTGG GAAACTGAAA 1560

GAGGAAAACT TCTTTAACCC CAAAGAAGAA GTGAAGCTGG AAGCGCATAT CAGAGTGC 1620  
 TCTTCCACAG CTGGCCGGGT GATTGGAAA GGTGGCAAGA CCGTGAACGA ACTGCAGAAC 1680  
 TTAACCAGTG CAGAAAGTCAT CGTGCCTCGT GACCAAACGC CAGATGAAA TGAGGAAGTG 1740  
 ATCGTCAGAA TTATCAGGGCA CCTCTTGCCT AGCCAGACTG CACAGCGAA GATCAGGGAA 1800  
 ATTGTACAAC AGGTGAAGCA GCAGGAGCAG AAATACCCCTC AGGGAGTCGC CTCACAGCGC 1860  
 AGCAAGTGAAG GCTCCCACAG GCACCAAGCAA AACAAACGGAT GAATGTAGCC CTTCCAACAC 1920  
 CTGACAGATA GAGACCAAAAC GCAGCCAGCC AGATCGGGAG CAAACCAAAG ACCATCTGAG 1980  
 GAATGAGAAG TCTGCGGAGG CGGCCAGGGGA CTCTGCCAG GCCCTGAGAA CCCCAGGGGC 2040  
 CGAGGAGGGGG CGGGGAAGGT CAGCCAGGTT TGCCAGAAC ACCGAGCCCC GCCTCCCGCC 2100  
 CCCCAGGGCT TCTGCAGGCT TCAGCCATCC ACTTCACCAT CCACCTCGGAT CTCTCCGTGAA 2160  
 CTCACACGAC GCTATCCCTT TAGTTGAAC TAACATAGGT GAACGTGTTA AAAGCCAAGC 2220  
 AAAATGCACA CCCTTTTCT GTGGCAAATC GTCTCTGTAC ATGTGTGTC ATATTAGAAA 2280  
 GGGGAAGATGT TAAGATATGT GGCCTGTGGG TTACACAGGG TCCCTGCAGC GGTAAATATAT 2340  
 TTTAGAAAATA ATATATCAA TAACTCACACT AACTCCAATT TTAATCAAT TATTAATTTT 2400  
 TTTTCTTTT TAAAGAGAAA GCAGGCTTT CTAGACTTTA AGAATAAAG TCTTGGGAG 2460  
 GTCTCACGGT GTAGAGAGGA GCTTGTAGGC CACCCGCACA AAATTCACCC AGAGGGAAAT 2520  
 CTCGTCGAA GGACACTCAC GGCAGTCTG GATCACCTGT GTATGTCAAC AGAAGGGATA 2580  
 CCGTCTCCTT GAAGAGGAAA CTCTGTCACT CCTCATGCCT GTCTAGCTCA TACACCCATT 2640  
 TCTCTTIGCT TCACAGGTT TAAACTGGTT TTTGTCATAC TGCTATATAA TTCTCTGTCT 2700  
 CTCTCTGTGTT ATCTCTCCCC TCCCTCCCT CCCCCTCTTC TCCATCTCCA TTCTTGTAA 2760  
 TTTCTCTCATC CCTCCATCTC AATCCCGTAT CTACGCCACCC CCCCCCCCCC AGGCAAAGCA 2820  
 GTGCTCTGAG TATCACATCA CACAAAAGGA ACAAAAGCGA AACACACAAA CCAGCCTCAA 2880  
 CTTACACTTG GTTACTCAAAGA AGAACAGAG TCAATGGTAC TTGTCTTAGC GTTTGGAAG 2940  
 AGGAAAACAG GAACCCACCA AACCAACCAA TCAACCAAAC AAAGAAAAAA TTCCACAAATG 3000  
 AAAGAATGTA TTTTGTCTTT TTGCAATTG GTGTATAAGC CATCAATATT CAGCAGGAAATG 3060  
 ATTCCCTTCTT TAAAAAAAATAA AAATGTGGAG GAAAGTAGAA ATTTACCAAG GTTGTGGCC 3120  
 CAGGGCGTTA AATTACACAGA TTTTTTAAAC GAGAAAACACAGAAGAA GCTACCTCAG 3180  
 GTGTTTTAC CTCAGCACCT TGCTCTTG TGCTCTTAG AGATTTGTA AAGCTGATAG 3240  
 TTGGAGCAATT TTTTATTTT TTTAATAAA ATGAGTTGGA AAAAAAATAA GATATCAACT 3300  
 GCCAGCGTGG AGAAGGTGAC AGTCCAAGTG TGCAACAGCT GTTCTGAATT GTCTCCGCT 3360  
 AGCCAAGAAC CNATATGGCC TTCTTTGGA CAAACCTTGA AAATGTTAT TT 3412

<210> 7  
 <211> 1946  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <400> 7

*C2*  
*W*  
 GCTGTAGCGG AGGGGCTGGG GGGCTGCTCT GTCCCTTCC TTGCGCGCTG CGGCCTCAGC 60  
 CCACCCAGAG GCCGGGGTGG GAGGGCGAGT GCTCAGCTTC CGGGGTTAGG AGCCGGAAAA 120  
 TTCAAATCCG AAATATTCCA CCCCAGCTCC GATGGGAAGT ACTGGACAGC CTGCTGGCTC 180  
 AGTATGGTAC AGTAGAGAAC TGTGAGCAAG TGAAACCCGA GAGTGGAGACG GCAGTGGTGA 240  
 ATGTCACCTA TTCCAACCGG GAGCAGACCA GGCAAGCCAT CATGAAGCTG AATGGCCACC 300  
 AGTTGGAGAA CCATGCCCTG AAGGTCTCT ACATCCCCGA TGAGCAGATA GCACAGGGAC 360  
 CTGAGAATGG GCGCCGAGGG GGCTTGGCT CTCGGGGTCA GCCCCGCCAG GGCTCACCTG 420  
 TGGCAGCGGG GCCCCCAGCC AAGCAGCAGC AAGTGGACAT CCCCCCTTCGG CTCCTGGTGC 480  
 CCACCCAGTA TGTTGGTGC ATTATTGGCA AGGAGGGGGC CACCATCCGC AACATCACAA 540  
 AACAGACCA GTCCAAGATA GACGTGCATA GGAAGGAGAA CGCAGGTGCA GCTGAAAAG 600  
 CCATCAGTGT GCACTCCACC CCTGAGGGCT GCTCCCTCCGC TTGTAAGATG ATCTTGGAGA 660  
 TTATGCATA AGAGGCTAAAG GACACAAAAA CGGCTGACGA GTTCCCCCTG AAGATCCTGG 720  
 CCCATAATAA CTTTGTAGGG CGTCTCATTC GCAAGGAAGG ACGGAACCTG AAGAAGGTAG 780  
 AGCAAGATAC CGAGACAAAAA ATCACCATCT CCTCGTTGCA AGACCTTACC CTTTACAACC 840  
 CTGAGAGGAC CATCACTGTG AAGGGGGGCCA TCGAGAATTG TTGCAAGGGCC GAGCAGGAAA 900  
 TAATGAAGAA AGTTCGGGAG GCCTATGAGA ATGATGTGGC TGCCATGAGC TCTCACCTGA 960  
 TCCCTGGCT GAACCTGGCT GCTGTAGGTC TTTTCCCGC TTCACTCCAGC GCAGTCCCGC 1020  
 CGCCTCCAG CAGCGTTACT GGGGCTGTC CCTATAGCTC CTTTATGCAAG GCTCCCGAGC 1080  
 AGGAGATGGT GCAGGTGTTT ATCCCCGCC AGGCAGTGGG CGCCATCATC GGCAAGAAGG 1140  
 GGCAGCACT CAAACAGCTC TCCCGGTTG CCAGCGCCTC CATCAAGATT GCACCAACCG 1200

AAACACCTGA CTCCAAAGTT CGTATGGTTA TCATCACTGG ACCGCCAGAG GCCCAATTCA 1260  
 AGGCTCAGGG AAGAACATAT GCCAAACCTCA AGGAGGAGAA CTCTTIGGT CCCAAGGAGG 1320  
 AAGTGAAGCT GGAGACCCAC ATACGTGTGC CAGCATCAGC AGCTGGCCCG GTCATTGGCA 1380  
 AAGGTGGAAA AACGGTGAC GAGTTGCAGA ATTTGACGGC AGCTGAGGTG GTAGTACCAA 1440  
 GAGACCAGAC CCCTGATGAG AACGACCAAGG TCATCGTGAA AATCATCGGA CATTCTATG 1500  
 CCAGTCAGAT GGCTAACCGG AAGATCCGAG ACATCCTGGC CCAGGTTAAG CAGCAGCATC 1560  
 AGAAGGGACA GAGTAACCAAG GCCCAGGCAC CGAGGAAGTG ACCAGCCCCCT CCCTGTCCCT 1620  
 TNGAGTCCAG GACAACAAACG GGCAGAAATC GAGAGTGTGC TCTCCCCGGC AGGCCTGAGA 1680  
 ATGAGTGGGA ATCCGGGACA CNTGGGCCGG GCTGTAGATC AGGTTTGCCT ACTTGATTGA 1740  
 GAAAGATGTT CCAGTGAGGA ACCCTGATCT NTCAGCCCCA AACACCCACC CAATTGGCCC 1800  
 AACACTGTNT GCCCCCTGGG GTGTCAGAAA TTNTAGCGCA AGGCACCTTT AAACGTGGAT 1860  
 TGTTAAAGA AGCTCTCCAG GCCCCACCAA GAGGGTGGAT CACACCTCAG TGGGAAGAAA 1920  
 AATAAAATT CCTTCAGGTT TTAAAAA 1946

<210> 8  
 <211> 3283  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <400> 8

C2  
 Cw

GGCAGCGGAG GAGGCAGGA GCGCCGGGT A CGGGGCCGGG GGAGCCGCGG GCTCTCGGGG 60  
 AAGAGACCGA TGATGAACAA GCTTTACATC GGGAAACCTGA GCCCCGCCGT CACCGCCGAC 120  
 GACCTCCGGC AGCTCTTGG GGACAGGAAG CTGCCCTGG CGGGACAGGT CCTGCTGAAG 180  
 TCCGGCTACG CCTCGTGGG CTACCCCGAC CAGAACTGGG CCATCCGCGC CATCGAGACC 240  
 CTCTCGGGTA AAGTGGATT GCATGGAAA ATCATGGAAG TTGATTACTC AGTCTCTAAA 300  
 AAGCTAAGGA GCAGGAAAAT TCAGATTGCA AACATCCCTC CTCACCTGCA GTGGGAGGTG 360  
 TTGGATGGAC TTTGGCTCA ATATGGACA GTGGAGAATG TGGAAACAAGT CAACACAGAC 420  
 ACAGAAACCG CCGTGTCAA CGTCACATAT GCAACAAAGAG AAGAAGCAAA AATAGCCATG 480  
 GAGAAGCTAA GCGGGCATCA GTTTGAGAAC TACTCCTTCA AGATTTCTTA CATCCGGAT 540  
 GAAGAGGTGA GCTCCCTTC GCCCCCTCAG CGAGGCCAGC GTGGGGACCA CTCTCCCGG 600  
 GAGCAAGGCC ACGCCCCCTGG GGGCACTTCT CAGGCCAGAC AGATTGATT CCCGCTGCGG 660  
 ATCCTGGTCC CCACCCAGTT GTTGGTGCC ATCATGGAA AGGAGGGCTT GACCATAAAG 720  
 AACATCACTA AGCAGACCCA GTCCCGGGT GATATCCATA GAAAAGAGAAA CTCTGGAGCT 780  
 GCAGAGAACGCTGTCACCAT CCATGCCACC CCAGAGGGGA CTTCTGAAGC ATGCCGCATG 840  
 ATTCTGAAA TCATGGAGAA AGAGGCAGAT GAGACCAAAC TAGCCGAAGA GATTCCCTG 900  
 AAAATCTGG CACACAATGG CTTGGTTGGA AGACTGATTG GAAAAGAAGG CAGAAATTG 960  
 AAGAAAATTG AACATGAAAC AGGGACCAAG ATAACAATCT CATCTTGCA GGATTTGAGC 1020  
 ATATACAACC CGGAAAGAAC CATCACTGTG AAGGGCACAG TTGAGGCCTG TGCCAGTGC 1080  
 GAGATAGAGA TTATGAAGAA GCTGCGTGAG GCCTTGAAA ATGATATGCT GGCTGTTAAC 1140  
 ACCCACTCCG GATACTTCTC CAGCTGTAC CCCCACATC ACCCTGGCCC GTTCCCGCAT 1200  
 CATCACTCTT ATCCAGAGCA GGAGATTGTG AATCTCTTCA TCCCAACCCA GGCTGTGGC 1260  
 GCCATCATCG GGAAGAAGGG GGCACACATC AAACAGCTGG CGAGATTGCG CGGAGCCTCT 1320  
 ATCAAGATTG CCCCTGCGGA AGGCCAGAC GTCAGCAGAA GGATGGTCT CATCACCGGG 1380  
 CCACCGGAAG CCCAGITCAA GGGCCAGGGA CGGAATCTTG GGAACATGAA AGAGGAAAC 1440  
 TTCTTTAACCC CCAAAGAAGA AGTGAAGCTG GAAGCGCATC TCAGAGTGCC CTCTTCCACA 1500  
 GCTGGCCGGG TGATGGCAA AGGTGGCAAG ACCGTGAACG AACTGCAGAA CTTAACCAAGT 1560  
 GCAGAAGTCA TCGTGCCTCG TGACCAAACG CCAGATGAAA ATGAGGAAGT GATCGTCAGA 1620  
 ATTATCGGGC ACTTCTTTGC TAGCCAGACT GCACAGCGCA AGATCAGGGAA ATTGTACAA 1680  
 CAGGTGAAGC AGCAGGAGCA GAAATACCT CAGGGAGTCG CCTCACAGCG CAGCAAGTGA 1740  
 GGCTCCCACA GGCACCCAGCA AAACAAACGA TGAATGTAGC CTTCCAACA CCTGACAGAA 1800  
 TGAGACCAAA CGCAGCCAGC CAGATGGGA GCAAACCAA GACCATCTGA GGAATGAGAA 1860  
 GTCTGCGGAG GCGGCCAGGG ACTCTGCCGA GGCCTTGAGA ACCCCAGGGG CCGAGGAGGG 1920  
 GCGGGGAAGG TCAGCCAGGT TTGCCAGAAC CACCGAGGCC CGCCTCCCGC CCCCCAGGGC 1980  
 TTCTGCAGGC TTCAGCCATC CACTTCACCA TCCACTCGGA TCTCTCTGA ACTCCCACGA 2040  
 CGCTATCCCT TTTAGTTGAA CTAACATAGG TGAACGTGTT CAAAGCCAAG CAAATGCAC 2100  
 ACCCTTTTC TGTGGCAAAT CGTCTCTGTA CATGTGTGTA CATATTAGAA AGGGAAAGATG 2160  
 TTAAGATATG TGGCCTGTGG GTTACACAGG GTGCCCTGCAG CGGTAATATA TTTTAGAAAT 2220  
 AATATATCAA ATAATCAAC TAACTCCAAT TTTTAATCAA TTATTAATT TTTTTCTTT 2280

TTAAAGAGAA AGCAGGCTTT TCTAGACTTT AAAGAATAAA GTCTTGGGA GGTCTCACGG 2340  
TGTAGAGAGG AGCTTGAGG CCACCCGCAC AAAATTCAAC CAGAGGGAAA TCTCGTCGGA 2400  
AGGACACTCA CGGCAGTCT GGATCACCTG TGTATGTCAA CAGAAGGGAT ACCGTCTCCT 2460  
TGAAGAGGAA ACTCTGTAC TCCATGCC TGTCTAGCTC ATACACCCAT TTCTCTTGC 2520  
TTCACAGGTT TTAAACTGGT TTTTGCATA CTGCTATATA ATTCTCTGTC TCTCTCTGTT 2580  
TATCTCTCCC CTCCCTCCCC TCCCTTCTT CTCCATCTCC ATTCTTTGA ATTTCTCAT 2640  
CCCTCCACT CAATCCCGTA TCTACGCACC CCCCCCCCCC CAGGCAAAGC AGTGCTCTGA 2700  
GTATCACATC ACACAAAAGG AACAAAAGCG AAACACACAA ACCAGCCTCA ACTTACACTT 2760  
GGTTACTCAA AAGAACAAAGA GTCAATGGTA CTTGTCCTAG CGTTTTGGAA GAGGAAAACA 2820  
GGAACCCACC AAACCAACCA ATCAACCAAA CAAAGAAAAA ATTCCACAAAT GAAAGAATGT 2880  
ATTTTGTCTT TTTGATTTT GGTGTATAAG CCATCAATAT TCAGCAAAT GATTCCCTTC 2940  
TTTAAAAAAA AAAATGTGGA GGAAAGTAGA AATTACCAA GTTGTGTCG GCAGGGCGTT 3000  
AAATTCAACAG ATTTTTTAA CGAGAAAAAC ACACAGAAGA AGCTACCTCA GGTGTTTTA 3060  
CCTCAGCACC TTGCTCTGT GTTCCCTTA GAGATTTGT AAAGCTGATA GTGGAGCAT 3120  
TTTTTTATT TTTTAATAAA AATGAGTTGG AAAAAAAATA AGATATCAAC TGCCAGCCTG 3180  
GAGAAGGTGA CAGTCCAAGT GTGCAACAGC TGTTCTGAAT TGCTTCCGC TAGCCAAGAA 3240  
CCNATATGGC CTTCTTTGG ACAAACCTTG AAAATGTTA TTT 3283

<210> 9  
<211> 21  
<212> DNA  
<213> Homo sapiens  
<220>  
<400> 9  
gaagtatct tcaaggacgc c 21

<210> 10  
<211> 22  
<212> DNA  
<213> Homo sapiens  
<220>  
<400> 10  
ctgcaagggg ttttgtggg cg 22

02  
w/  
<210> 11  
<211> 22  
<212> DNA  
<213> Homo sapiens  
<220>  
<400> 11  
tccttgcgcg ctgcggcctc ag 22

<210> 12  
<211> 23  
<212> DNA  
<213> Homo sapiens  
<220>  
<400> 12  
ccaaactggc gccattcagc ttc 23

<210> 13  
<211> 22  
<212> DNA  
<213> Homo sapiens  
<220>  
<400> 13  
gcttttggg gacaggaagg tc 22

<210> 14  
<211> 22

<212> DNA  
<213> Homo sapiens  
<220>  
<400> 14  
gacgttgaca acggcggtt ct 22

<210> 15  
<211> 12  
<212> DNA  
<213> artificial sequence  
<220>  
<221> adaptor  
<222> 1...12  
<400> 15  
aatttgcgtt ga 12

<210> 16  
<211> 12  
<212> DNA  
<213> artificial sequence  
<220>  
<221> adaptor  
<222> 1...12  
<400> 16  
aatttgcgttca tg 12

<210> 17  
<211> 12  
<212> DNA  
<213> artificial sequence  
<220>  
<221> adaptor  
<222> 1...12  
<400> 17  
aattttccct cg 12

C2  
*Ca*  
<210> 18  
<211> 24  
<212> DNA  
<213> artificial sequence  
<220>  
<221> adaptor  
<222> 1...24  
<400> 18  
agcactctcc agcccttcac catg 24

<210> 19  
<211> 24  
<212> DNA  
<213> artificial sequence  
<220>  
<221> adaptor  
<222> 1...24  
<400> 19  
accgacgtcg actatcatg catg 24

<210> 20  
<211> 24

<212> DNA  
<213> artificial sequence  
<220>  
<221> adaptor  
<222> 1...24  
<400> 20

aggcaactgt gctatccgag catg 24

<210> 21  
<211> 8  
<212> DNA  
<213> artificial sequence  
<220>  
<221> adaptor  
<222> 1...8  
<400> 21

gtgagagg 8

<210> 22  
<211> 8  
<212> DNA  
<213> artificial sequence  
<220>  
<221> adaptor  
<222> 1...8  
<400> 22

catggatg 8

C2  
2  
<210> 23  
<211> 8  
<212> DNA  
<213> artificial sequence  
<220>  
<221> adaptor  
<222> 1...8  
<400> 23

cctcggtata 8